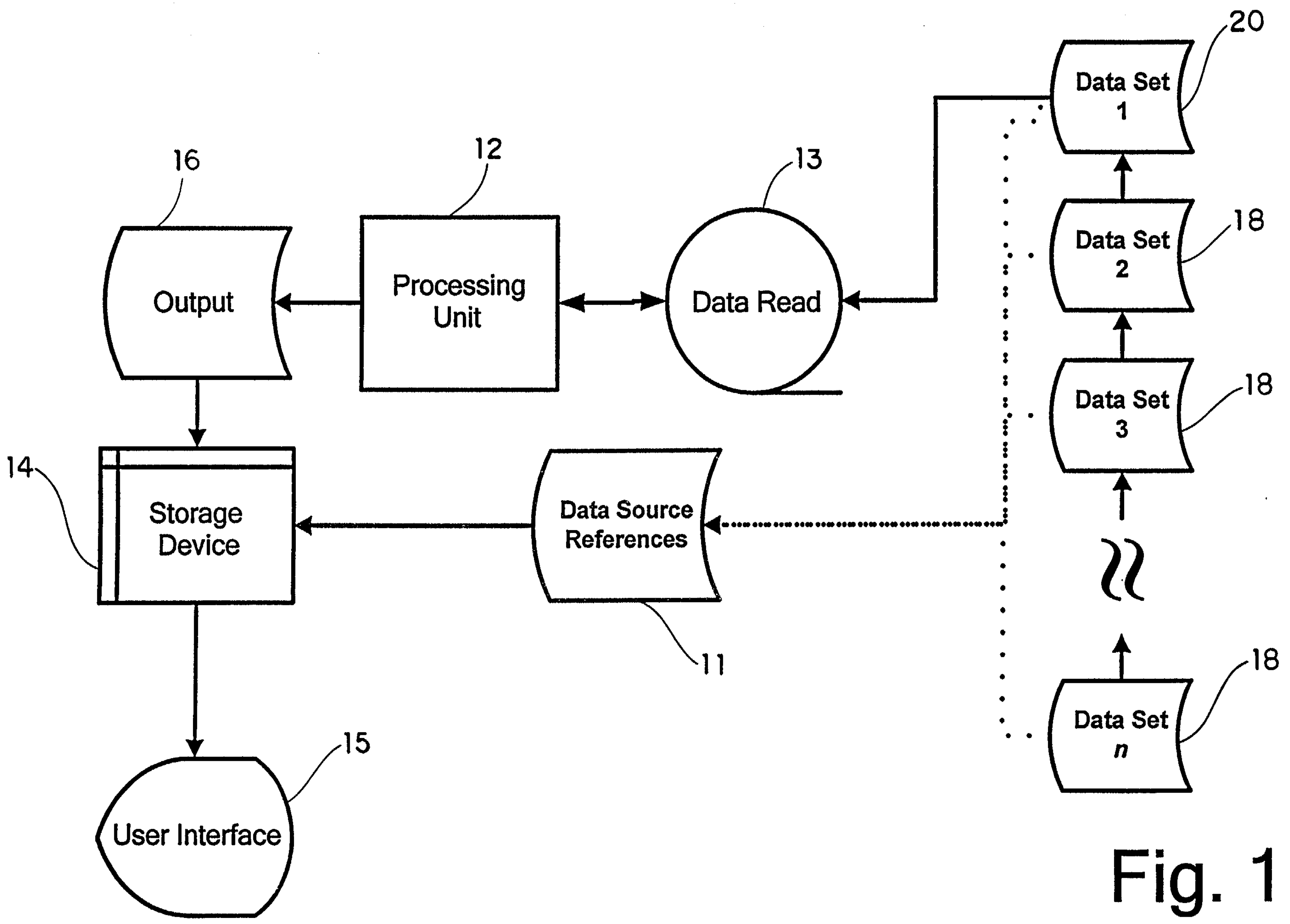


20
Data Set 1
18
Data Set 2
18
Data Set 3
18
Data Set n
13
Data Read
12
Processing Unit
16
Output
11
Data Source References
14
Storage Device
15
User Interface

Fig. 1

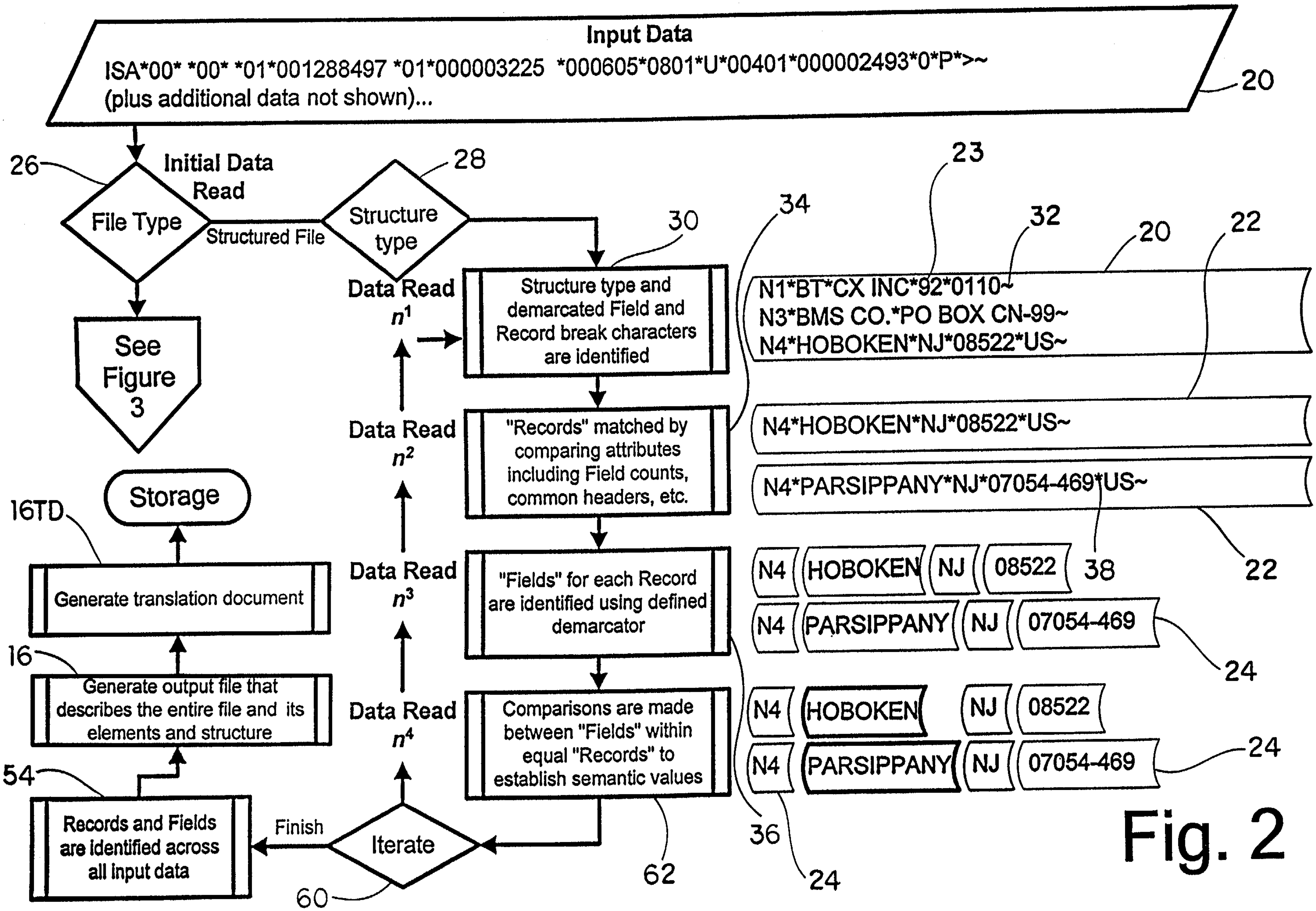

Input Data
ISA*00* *00* *01*001288497 *01*000003225 *000605*0801*U*00401*000002493*0*P*>~
(plus additional data not shown)...
20
26
Initial Data Read
File Type
Structured File
See Figure 3
28
Structure type
Data Read n^1
30
Structure type and demarcated Field and Record break characters are identified
Data Read n^2
34
"Records" matched by comparing attributes including Field counts, common headers, etc.
Data Read n^3
"Fields" for each Record are identified using defined demarcator
Data Read n^4
Comparisons are made between "Fields" within equal "Records" to establish semantic values
62
Iterate
60
Finish
Records and Fields are identified across all input data
54
Generate output file that describes the entire file and its elements and structure
16
Generate translation document
16TD
Storage
23
32
20
22
N1*BT*CX INC*92*0110~
N3*BMS CO.*PO BOX CN-99~
N4*HOBOKEN*NJ*08522*US~
N4*HOBOKEN*NJ*08522*US~
N4*PARSIPPANY*NJ*07054-469*US~
22
38
N4
HOBOKEN
NJ
08522
N4
PARSIPPANY
NJ
07054-469
24
N4
HOBOKEN
NJ
08522
N4
PARSIPPANY
NJ
07054-469
24
36
24

Fig. 2

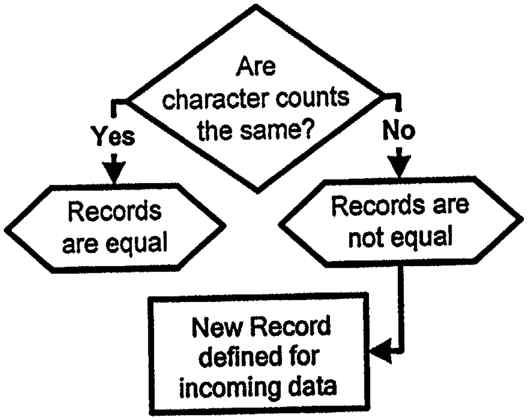

Are
character counts
the same?
Yes
No
Records
are equal
Records are
not equal
New Record
defined for
incoming data

Fig. 2A

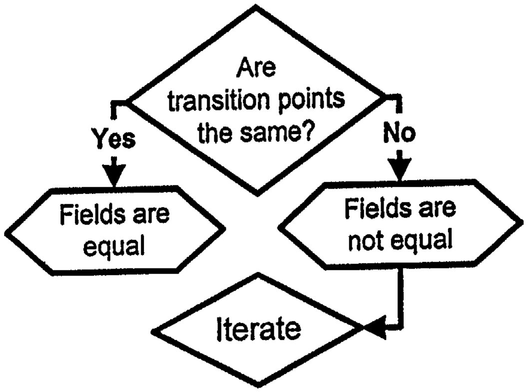

Are
transition points
the same?
Yes
No
Fields are
equal
Fields are
not equal
Iterate

Fig. 2B

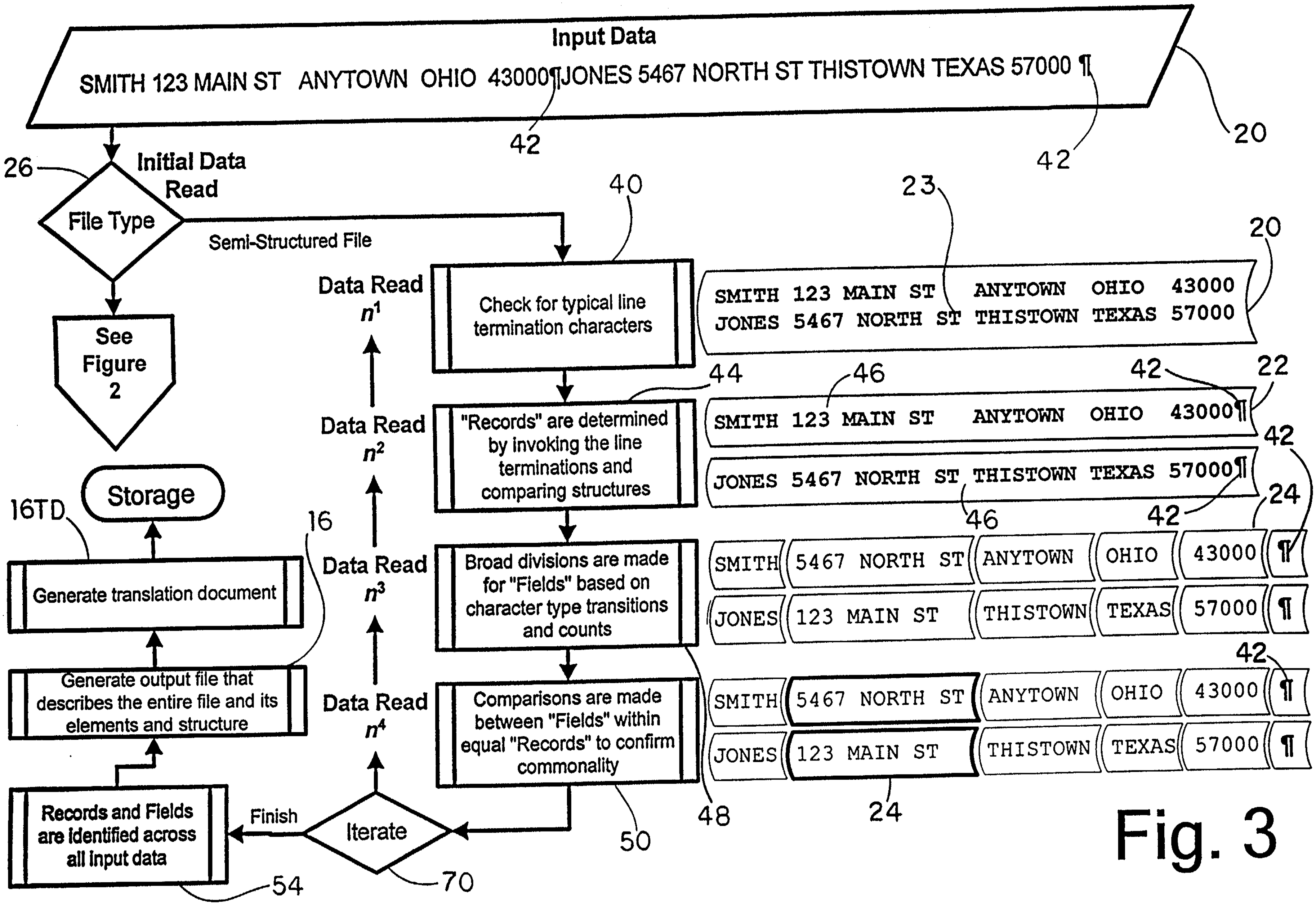
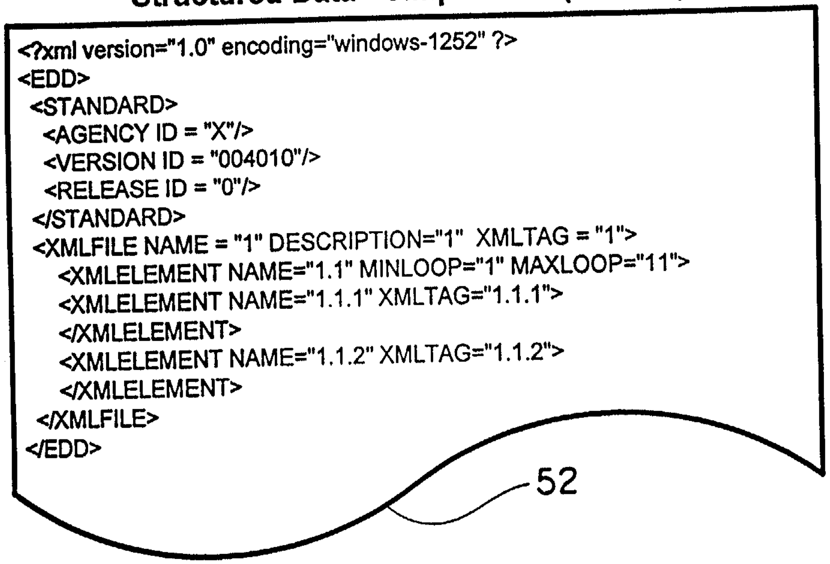

Input Data
SMITH 123 MAIN ST ANYTOWN OHIO 43000¶JONES 5467 NORTH ST THISTOWN TEXAS 57000 ¶
42
42
20
26
Initial Data Read
File Type
Semi-Structured File
See Figure 2
Storage
16TD
Generate translation document
16
Generate output file that describes the entire file and its elements and structure
Records and Fields are identified across all input data
54
Finish
Iterate
70
Data Read n^1
Data Read n^2
Data Read n^3
Data Read n^4
40
Check for typical line termination characters
44
"Records" are determined by invoking the line terminations and comparing structures
Broad divisions are made for "Fields" based on character type transitions and counts
Comparisons are made between "Fields" within equal "Records" to confirm commonality
50
48
23
20
SMITH 123 MAIN ST ANYTOWN OHIO 43000
JONES 5467 NORTH ST THISTOWN TEXAS 57000
46
42
22
SMITH 123 MAIN ST ANYTOWN OHIO 43000¶
JONES 5467 NORTH ST THISTOWN TEXAS 57000¶
42
46
42
24
SMITH 5467 NORTH ST ANYTOWN OHIO 43000 ¶
JONES 123 MAIN ST THISTOWN TEXAS 57000 ¶
42
SMITH 5467 NORTH ST ANYTOWN OHIO 43000 ¶
JONES 123 MAIN ST THISTOWN TEXAS 57000 ¶
24

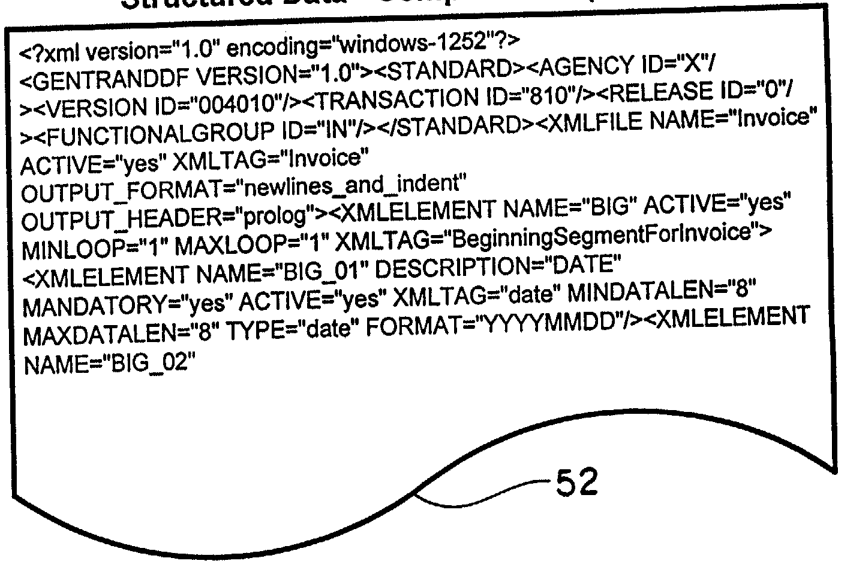
Fig. 3

Structured Data - Simple XML (Portion)



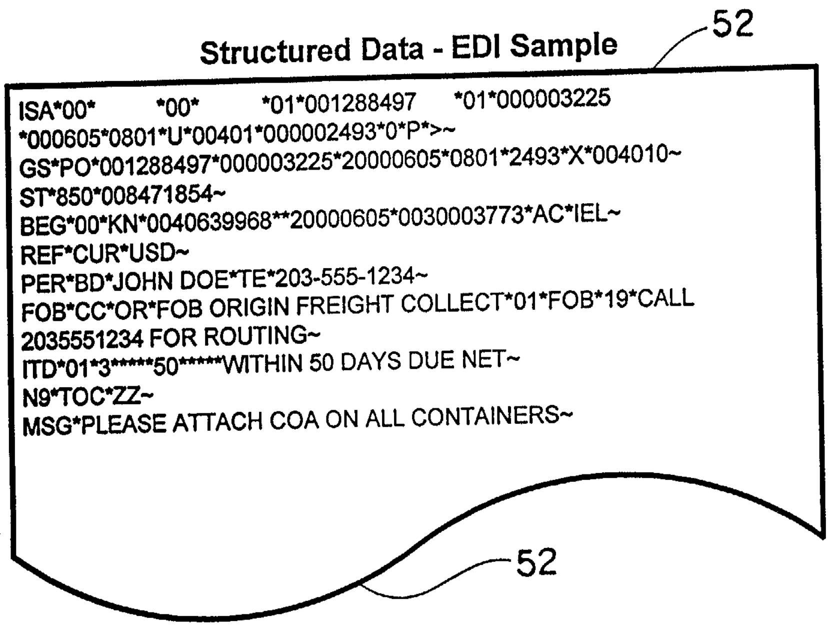
```
<?xml version="1.0" encoding="windows-1252" ?>
<EDD>
 <STANDARD>
  <AGENCY ID = "X"/>
  <VERSION ID = "004010"/>
  <RELEASE ID = "0"/>
 </STANDARD>
 <XMLFILE NAME = "1" DESCRIPTION="1"  XMLTAG = "1">
   <XMLELEMENT NAME="1.1" MINLOOP="1" MAXLOOP="11">
   <XMLELEMENT NAME="1.1.1" XMLTAG="1.1.1">
   </XMLELEMENT>
   <XMLELEMENT NAME="1.1.2" XMLTAG="1.1.2">
   </XMLELEMENT>
 </XMLFILE>
</EDD>
```

Structured Data - Complex XML (Portion)



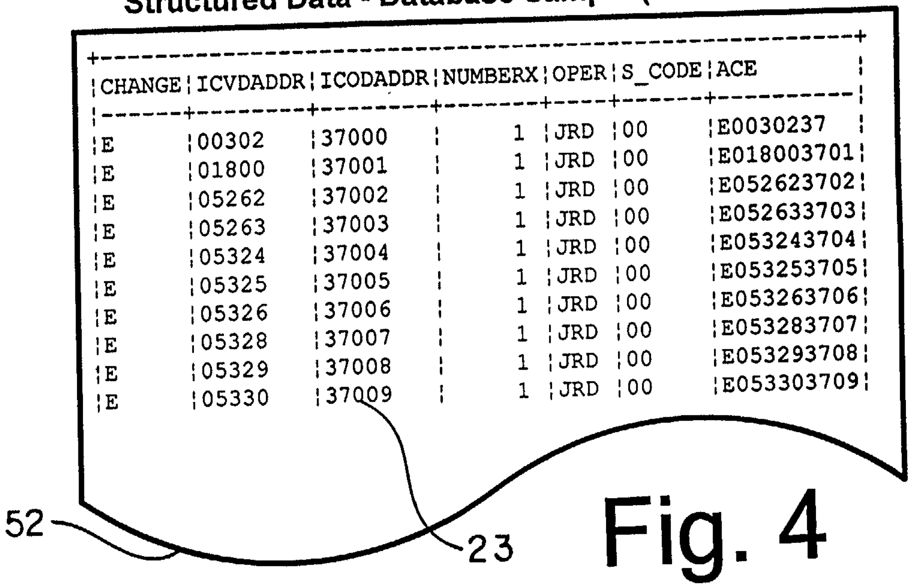
```
<?xml version="1.0" encoding="windows-1252"?>
<GENTRANDDF VERSION="1.0"><STANDARD><AGENCY ID="X"/
><VERSION ID="004010"/><TRANSACTION ID="810"/><RELEASE ID="0"/
><FUNCTIONALGROUP ID="IN"/></STANDARD><XMLFILE NAME="Invoice"
ACTIVE="yes" XMLTAG="Invoice"
OUTPUT_FORMAT="newlines_and_indent"
OUTPUT_HEADER="prolog"><XMLELEMENT NAME="BIG" ACTIVE="yes"
MINLOOP="1" MAXLOOP="1" XMLTAG="BeginningSegmentForInvoice">
<XMLELEMENT NAME="BIG_01" DESCRIPTION="DATE"
MANDATORY="yes" ACTIVE="yes" XMLTAG="date" MINDATALEN="8"
MAXDATALEN="8" TYPE="date" FORMAT="YYYYMMDD"/><XMLELEMENT
NAME="BIG_02"
```

Structured Data - EDI Sample



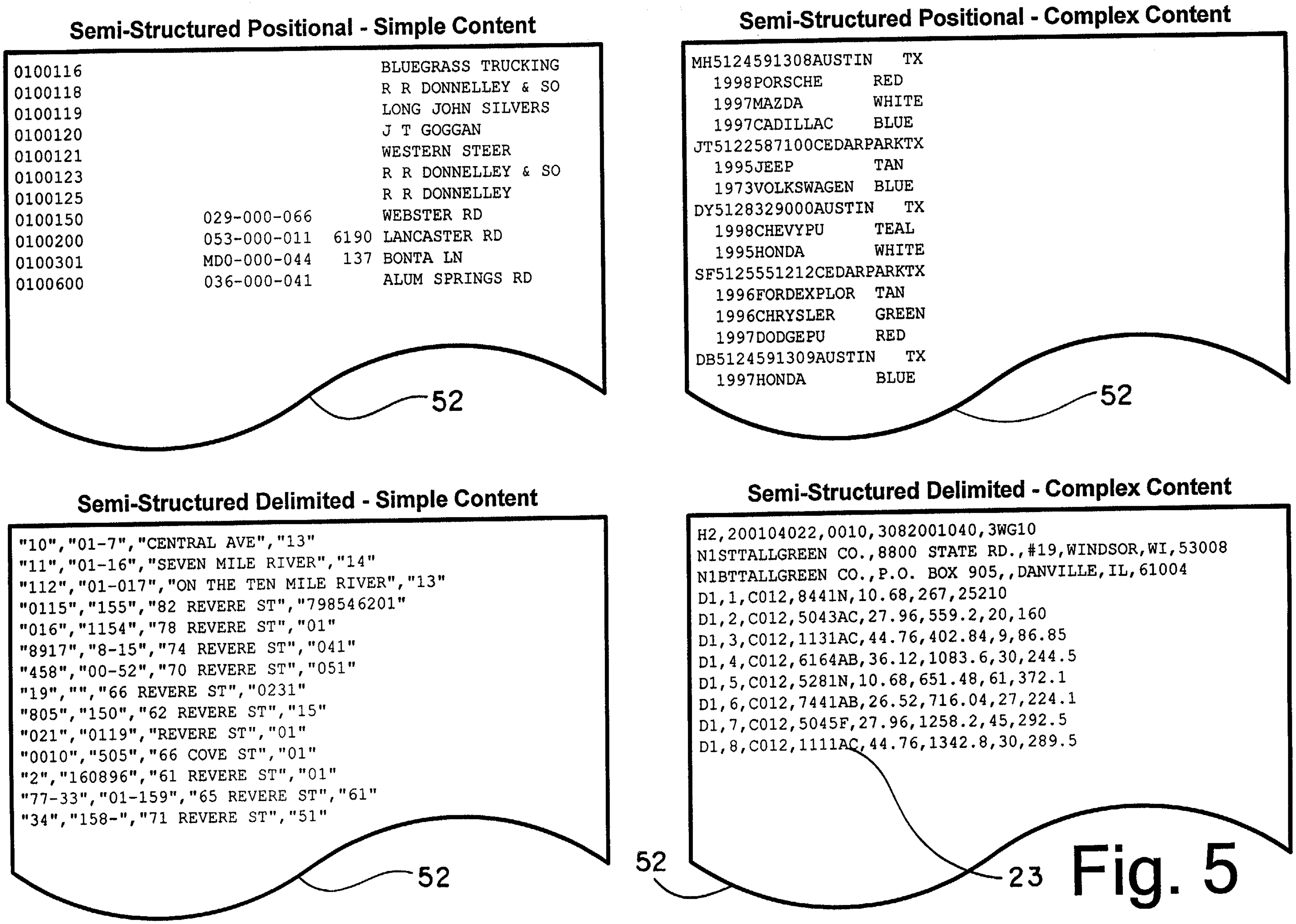
```
ISA*00*        *00*        *01*001288497     *01*000003225
*000605*0801*U*00401*000002493*0*P*>~
GS*PO*001288497*000003225*20000605*0801*2493*X*004010~
ST*850*008471854~
BEG*00*KN*0040639968**20000605*0030003773*AC*IEL~
REF*CUR*USD~
PER*BD*JOHN DOE*TE*203-555-1234~
FOB*CC*OR*FOB ORIGIN FREIGHT COLLECT*01*FOB*19*CALL
2035551234 FOR ROUTING~
ITD*01*3*****50*****WITHIN 50 DAYS DUE NET~
N9*TOC*ZZ~
MSG*PLEASE ATTACH COA ON ALL CONTAINERS~
```

Structured Data - Database Sample (Browse view)



```
+--------------------------------------------------------------------+
|CHANGE|ICVDADDR|ICODADDR|NUMBERX|OPER|S_CODE|ACE                    |
|------+--------+--------+-------+----+------+-----------------------|
|E     |00302   |37000   |      1|JRD |00    |E0030237               |
|E     |01800   |37001   |      1|JRD |00    |E018003701|
|E     |05262   |37002   |      1|JRD |00    |E052623702|
|E     |05263   |37003   |      1|JRD |00    |E052633703|
|E     |05324   |37004   |      1|JRD |00    |E053243704|
|E     |05325   |37005   |      1|JRD |00    |E053253705|
|E     |05326   |37006   |      1|JRD |00    |E053263706|
|E     |05328   |37007   |      1|JRD |00    |E053283707|
|E     |05329   |37008   |      1|JRD |00    |E053293708|
|E     |05330   |37009   |      1|JRD |00    |E053303709|
```

Fig. 4



Semi-Structured Positional - Simple Content

```
0100116                                  BLUEGRASS TRUCKING
0100118                                  R R DONNELLEY & SO
0100119                                  LONG JOHN SILVERS
0100120                                  J T GOGGAN
0100121                                  WESTERN STEER
0100123                                  R R DONNELLEY & SO
0100125                                  R R DONNELLEY
0100150            029-000-066           WEBSTER RD
0100200            053-000-011     6190  LANCASTER RD
0100301            MD0-000-044      137  BONTA LN
0100600            036-000-041           ALUM SPRINGS RD
```

Semi-Structured Positional - Complex Content

```
MH5124591308AUSTIN    TX
  1998PORSCHE      RED
  1997MAZDA        WHITE
  1997CADILLAC     BLUE
JT5122587100CEDARPARKTX
  1995JEEP         TAN
  1973VOLKSWAGEN   BLUE
DY5128329000AUSTIN    TX
  1998CHEVYPU      TEAL
  1995HONDA        WHITE
SF5125551212CEDARPARKTX
  1996FORDEXPLOR   TAN
  1996CHRYSLER     GREEN
  1997DODGEPU      RED
DB5124591309AUSTIN    TX
  1997HONDA        BLUE
```

Semi-Structured Delimited - Simple Content

```
"10","01-7","CENTRAL AVE","13"
"11","01-16","SEVEN MILE RIVER","14"
"112","01-017","ON THE TEN MILE RIVER","13"
"0115","155","82 REVERE ST","798546201"
"016","1154","78 REVERE ST","01"
"8917","8-15","74 REVERE ST","041"
"458","00-52","70 REVERE ST","051"
"19","","66 REVERE ST","0231"
"805","150","62 REVERE ST","15"
"021","0119","REVERE ST","01"
"0010","505","66 COVE ST","01"
"2","160896","61 REVERE ST","01"
"77-33","01-159","65 REVERE ST","61"
"34","158-","71 REVERE ST","51"
```

Semi-Structured Delimited - Complex Content

```
H2,200104022,0010,3082001040,3WG10
N1STTALLGREEN CO.,8800 STATE RD.,#19,WINDSOR,WI,53008
N1BTTALLGREEN CO.,P.O. BOX 905,,DANVILLE,IL,61004
D1,1,C012,8441N,10.68,267,25210
D1,2,C012,5043AC,27.96,559.2,20,160
D1,3,C012,1131AC,44.76,402.84,9,86.85
D1,4,C012,6164AB,36.12,1083.6,30,244.5
D1,5,C012,5281N,10.68,651.48,61,372.1
D1,6,C012,7441AB,26.52,716.04,27,224.1
D1,7,C012,5045F,27.96,1258.2,45,292.5
D1,8,C012,1111AC,44.76,1342.8,30,289.5
```

Fig. 5

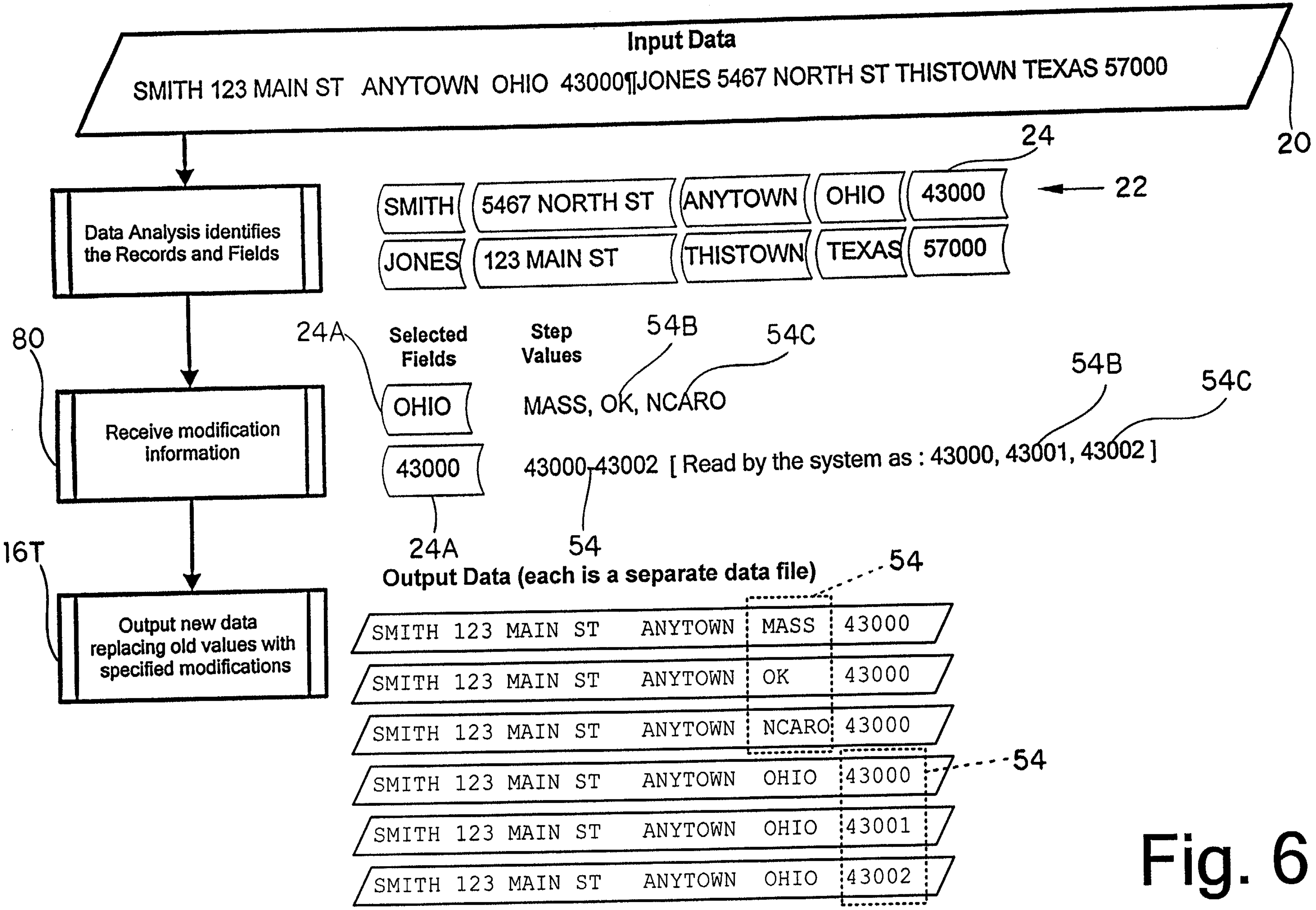

Input Data
SMITH 123 MAIN ST ANYTOWN OHIO 43000¶JONES 5467 NORTH ST THISTOWN TEXAS 57000
20
Data Analysis identifies the Records and Fields
24
SMITH
5467 NORTH ST
ANYTOWN
OHIO
43000
22
JONES
123 MAIN ST
THISTOWN
TEXAS
57000
80
Receive modification information
24A
Selected Fields
OHIO
43000
24A
Step Values
54B
54C
MASS, OK, NCARO
54B
54C
43000-43002 [Read by the system as : 43000, 43001, 43002]
54
16T
Output new data replacing old values with specified modifications
Output Data (each is a separate data file)
54
SMITH 123 MAIN ST ANYTOWN MASS 43000
SMITH 123 MAIN ST ANYTOWN OK 43000
SMITH 123 MAIN ST ANYTOWN NCARO 43000
SMITH 123 MAIN ST ANYTOWN OHIO 43000
54
SMITH 123 MAIN ST ANYTOWN OHIO 43001
SMITH 123 MAIN ST ANYTOWN OHIO 43002

Fig. 6